09/90
070

OIPE

ENTERED

RAW SEQUENCE LISTING DATE: 07/19/2002
PATENT APPLICATION: US/09/966,880A TIME: 13:48:23

Input Set : A:\06501-088001.txt
Output Set: N:\CRF3\07192002\I966880A.raw

```
 4 <110> APPLICANT: Honjo, Tasuku
 5       Muramatsu, Masamichi
 7 <120> TITLE OF INVENTION: NOVEL CYTIDINE DEAMINASE
10 <130> FILE REFERENCE: 06501-088001
12 <140> CURRENT APPLICATION NUMBER: 09/966,880A
13 <141> CURRENT FILING DATE: 2001-09-28
15 <150> PRIOR APPLICATION NUMBER: PCT/JP00/01918
16 <151> PRIOR FILING DATE: 2000-03-28
18 <150> PRIOR APPLICATION NUMBER: JP 11-371382
19 <151> PRIOR FILING DATE: 1999-12-27
21 <150> PRIOR APPLICATION NUMBER: JP 11-178999
22 <151> PRIOR FILING DATE: 1999-06-24
24 <150> PRIOR APPLICATION NUMBER: JP 11-87192
25 <151> PRIOR FILING DATE: 1999-03-29
27 <160> NUMBER OF SEQ ID NOS: 36
29 <170> SOFTWARE: FastSEQ for Windows Version 4.0
31 <210> SEQ ID NO: 1
32 <211> LENGTH: 2440
33 <212> TYPE: DNA
34 <213> ORGANISM: Mus musculus
36 <220> FEATURE:
37 <221> NAME/KEY: CDS
38 <222> LOCATION: (93)...(686)
40 <221> NAME/KEY: 5'UTR
41 <222> LOCATION: (1)...(92)
W--> 43 <221> 3'UTR
44 <222> LOCATION: (690)...(2440)
W--> 46 <221> misc_feature
47 <222> LOCATION: (1)...(2440)
48 <223> OTHER INFORMATION: n = A,T,C or G
W--> 50 <400> 1
51 ggcacgagca gcactgaagc agccttgctt gaagcaagct tcctttggcc taagactttg      60
52 agggagtcaa gaaagtcacg ctggagaccg at atg gac agc ctt ctg atg aag      113
53                                     Met Asp Ser Leu Leu Met Lys
54                                      1               5
56 caa aag aag ttt ctt tac cat ttc aaa aat gtc cgc tgg gcc aag gga      161
57 Gln Lys Lys Phe Leu Tyr His Phe Lys Asn Val Arg Trp Ala Lys Gly
58          10                  15                  20
60 cgg cat gag acc tac ctc tgc tac gtg gtg aag agg aga gat agt gcc      209
61 Arg His Glu Thr Tyr Leu Cys Tyr Val Val Lys Arg Arg Asp Ser Ala
62      25                  30                  35
64 acc tcc tgc tca ctg gac ttc ggc cac ctt cgc aac aag tct ggc tgc      257
65 Thr Ser Cys Ser Leu Asp Phe Gly His Leu Arg Asn Lys Ser Gly Cys
```

RAW SEQUENCE LISTING DATE: 07/19/2002
PATENT APPLICATION: US/09/966,880A TIME: 13:48:23

Input Set : A:\06501-088001.txt
Output Set: N:\CRF3\07192002\I966880A.raw

```
66  40                  45                  50                  55
68 cac gtg gaa ttg ttg ttc cta cgc tac atc tca gac tgg gac ctg gac      305
69 His Val Glu Leu Leu Phe Leu Arg Tyr Ile Ser Asp Trp Asp Leu Asp
70                   60                  65                  70
72 ccg ggc cgg tgt tac cgc gtc acc tgg ttc acc tcc tgg agc ccg tgc      353
73 Pro Gly Arg Cys Tyr Arg Val Thr Trp Phe Thr Ser Trp Ser Pro Cys
74              75                  80                  85
76 tat gac tgt gcc cgg cac gtg gct gag ttt ctg aga tgg aac cct aac      401
77 Tyr Asp Cys Ala Arg His Val Ala Glu Phe Leu Arg Trp Asn Pro Asn
78          90                  95                  100
80 ctc agc ctg agg att ttc acc gcg cgc ctc tac ttc tgt gaa gac cgc      449
81 Leu Ser Leu Arg Ile Phe Thr Ala Arg Leu Tyr Phe Cys Glu Asp Arg
82     105                 110                 115
84 aag gct gag cct gag ggg ctg cgg aga ctg cac cgc gct ggg gtc cag      497
85 Lys Ala Glu Pro Glu Gly Leu Arg Arg Leu His Arg Ala Gly Val Gln
86 120                 125                 130                 135
88 atc ggg atc atg acc ttc aaa gac tat ttt tac tgc tgg aat aca ttt      545
89 Ile Gly Ile Met Thr Phe Lys Asp Tyr Phe Tyr Cys Trp Asn Thr Phe
90                  140                 145                 150
92 gta gaa aat cgt gaa aga act ttc aaa gcc tgg gaa ggg cta cat gaa      593
93 Val Glu Asn Arg Glu Arg Thr Phe Lys Ala Trp Glu Gly Leu His Glu
94              155                 160                 165
96 aat tct gtc cgg cta acc aga caa ctt cgg cgc atc ctt ttg ccc ttg      641
97 Asn Ser Val Arg Leu Thr Arg Gln Leu Arg Arg Ile Leu Leu Pro Leu
98          170                 175                 180
100 tac gaa gtc gat gac ttg cga gat gca ttt cgt atg ttg gga ttt          686
101 Tyr Glu Val Asp Asp Leu Arg Asp Ala Phe Arg Met Leu Gly Phe
102     185                 190                 195
104 tgaaagcaac ctcctggaat gtcacacgtg atgaaatttc tctgaagaga ctggatagaa   746
105 aaacaaccct tcaactacat gtttttcttc ttaagtactc acttttataa gtgtaggggg   806
106 aaattatatg acttttaaa aaatacttga gctgcacagg accgccagag caatgatgta   866
107 actgagcttg ctgtgcaaca tcgccatcta ctggggaaca gcataacttc cagactttgg   926
108 gtcgtgaatg atgctctttt ttttcaacag catggaaaag catatggaga cgaccacaca   986
109 gtttgttaca cccaccctgt gttccttgat tcatttgaat tctcaggggt atcagtgacg  1046
110 gattcttcta ttctttccct ctaaggctca ctttcagggg tccttttctg acaaggtcac  1106
111 ggggctgtcc tacagtctct gtctgagcaa tcacaagcca ttctctcaaa aacattaata  1166
112 ctcaggcaca tgctgtatgt tttcactgtc cgtcgtgttt ttcacatttg tatgtgaaag  1226
113 ggcttggggt gggatttgaa gaatgcacga tcgcctctgg gtgatttcaa taaaggatct  1286
114 taaaatgcag atgaggacta cgaagaaatc actctgaaaa tgagttcacg cctcaagaag  1346
115 caaatcccct ggaaacacag actctttttc atttttaatg tcattagttt actcacagtc  1406
116 ttatcaagaa gaagagttca agggttcaac ccaattttca gatcgcgtcc cttaaacatc  1466
117 agtaattctg ttaaagggat caaacatcct tatttcttaa ctaactggtg ccttgctgta  1526
118 gagaaaggag caaagcgccc agatccaaag tatatagtta tcatagccag gaaccgctac  1586
119 tcgttttcca ttacaaatgg caaattcttc cccgggctct cctcatagtg cctgagacgg  1646
120 accacggagg tgatgaacct ccggattctc tggcccaaca cggtggaagc tctgcaaggg  1706
121 cgcagagaca gaatgcggca gaaattgccc ccgagtccca actctccttt ccttgcgacc  1766
122 ttgggaacaa gacttaaagg agcctgtgac ttagaaactt ctagtaatgg gtacctggga  1826
123 gtcgtttgag tatggggcag tgatttattc tctgtgatgg atgccaacac ggttaaacag  1886
124 aatttttagt ttttatatgt gtgtgatgct gctcccccaa attgttaact gtgtaagagg  1946
```

RAW SEQUENCE LISTING DATE: 07/19/2002
PATENT APPLICATION: US/09/966,880A TIME: 13:48:23

Input Set : A:\06501-088001.txt
Output Set: N:\CRF3\07192002\I966880A.raw

```
    125 gtggcaaaat agggaaagtg gcattcacct atagttccag cattcaggaa gctgaggcag   2006
    126 gaggattgta aatttgaggc cagtctgagc tgtaaggtga gaccctattt caaacaacac   2066
    127 agccagaatt gggttctggt aaatcatact taacaaggga aaaatgcaag acgcaagacc   2126
    128 gtggcaagga aatgacgctt tgcccaacga aatgtaggaa accaacatag actcccagtt   2186
    129 tgtccctctt tatgtctggt ctccctaaca acgatctttg ctaatgagaa aaatattaga   2246
    130 aaaaaatatc cctgtgcaat tatcacccag tcgccattat aatgcaatta aaaggcccac   2306
    131 aagaaatcct gtatacacga ccgttattta ttgtatgtaa gttgctgagg aagaggagaa   2366
W--> 132 aaaaataaag atcatccatt ccttcctgca aaaaaaaaaa aaaaaanaaa aaaaaaaaaa   2426
    133 aaaaaaaaaa aaaa                                                   2440
    135 <210> SEQ ID NO: 2
    136 <211> LENGTH: 198
    137 <212> TYPE: PRT
    138 <213> ORGANISM: Mus musculus
    140 <400> SEQUENCE: 2
    141 Met Asp Ser Leu Leu Met Lys Gln Lys Lys Phe Leu Tyr His Phe Lys
    142  1               5                  10                  15
    143 Asn Val Arg Trp Ala Lys Gly Arg His Glu Thr Tyr Leu Cys Tyr Val
    144             20                  25                  30
    145 Val Lys Arg Arg Asp Ser Ala Thr Ser Cys Ser Leu Asp Phe Gly His
    146         35                  40                  45
    147 Leu Arg Asn Lys Ser Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr
    148     50                  55                  60
    149 Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp
    150 65                  70                  75                  80
    151 Phe Thr Ser Trp Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Glu
    152                 85                  90                  95
    153 Phe Leu Arg Trp Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg
    154             100                 105                 110
    155 Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg
    156         115                 120                 125
    157 Leu His Arg Ala Gly Val Gln Ile Gly Ile Met Thr Phe Lys Asp Tyr
    158     130                 135                 140
    159 Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn Arg Glu Arg Thr Phe Lys
    160 145                 150                 155                 160
    161 Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Thr Arg Gln Leu
    162                 165                 170                 175
    163 Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp Asp Leu Arg Asp Ala
    164             180                 185                 190
    165 Phe Arg Met Leu Gly Phe
    166         195
    168 <210> SEQ ID NO: 3
    169 <211> LENGTH: 30
    170 <212> TYPE: DNA
    171 <213> ORGANISM: Artificial Sequence
    173 <220> FEATURE:
    174 <223> OTHER INFORMATION: Artificially synthesized primer sequence, AID138
    178 <400> SEQUENCE: 3
    179 ggaattcgcc atggacagcc ttctgatgaa                                    30
    181 <210> SEQ ID NO: 4
```

RAW SEQUENCE LISTING DATE: 07/19/2002
PATENT APPLICATION: US/09/966,880A TIME: 13:48:23

Input Set : A:\06501-088001.txt
Output Set: N:\CRF3\07192002\I966880A.raw

```
182 <211> LENGTH: 30
183 <212> TYPE: DNA
184 <213> ORGANISM: Artificial Sequence
186 <220> FEATURE:
187 <223> OTHER INFORMATION: Artificially synthesized primer sequence, AID161
191 <400> SEQUENCE: 4
192 gccgctcgag tcaaaatccc aacatacgaa                                        30
194 <210> SEQ ID NO: 5
195 <211> LENGTH: 25
196 <212> TYPE: DNA
197 <213> ORGANISM: Artificial Sequence
199 <220> FEATURE:
200 <223> OTHER INFORMATION: Artificially synthesized primer sequence, AID118
204 <400> SEQUENCE: 5
205 ggctgaggtt agggttccat ctcag                                             25
207 <210> SEQ ID NO: 6
208 <211> LENGTH: 25
209 <212> TYPE: DNA
210 <213> ORGANISM: Artificial Sequence
212 <220> FEATURE:
213 <223> OTHER INFORMATION: Artificially synthesized primer sequence, AID119
217 <400> SEQUENCE: 6
218 gagggagtca agaaagtcac gctgg                                            25
220 <210> SEQ ID NO: 7
221 <211> LENGTH: 2818
222 <212> TYPE: DNA
223 <213> ORGANISM: Homo sapiens
225 <220> FEATURE:
226 <221> NAME/KEY: CDS
227 <222> LOCATION: (80)...(673)
229 <221> NAME/KEY: 5'UTR
230 <222> LOCATION: (1)...(79)
W--> 232 <221> 3'UTR
233 <222> LOCATION: (677)...(2818)
W--> 235 <400> 7
236 agagaaccat cattaattga agtgagattt ttctggcctg agacttgcag ggaggcaaga     60
237 agacactctg gacaccact atg gac agc ctc ttg atg aac cgg agg aag ttt     112
238                     Met Asp Ser Leu Leu Met Asn Arg Arg Lys Phe
239                      1               5                  10
241 ctt tac caa ttc aaa aat gtc cgc tgg gct aag ggt cgg cgt gag acc      160
242 Leu Tyr Gln Phe Lys Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr
243             15                  20                  25
245 tac ctg tgc tac gta gtg aag agg cgt gac agt gct aca tcc ttt tca      208
246 Tyr Leu Cys Tyr Val Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser
247          30                  35                  40
249 ctg gac ttt ggt tat ctt cgc aat aag aac ggc tgc cac gtg gaa ttg      256
250 Leu Asp Phe Gly Tyr Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu
251      45                  50                  55
253 ctc ttc ctc cgc tac atc tcg gac tgg gac cta gac cct ggc cgc tgc      304
```

RAW SEQUENCE LISTING DATE: 07/19/2002
PATENT APPLICATION: US/09/966,880A TIME: 13:48:23

Input Set : A:\06501-088001.txt
Output Set: N:\CRF3\07192002\I966880A.raw

```
254 Leu Phe Leu Arg Tyr Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys
255  60                  65                  70                  75
257 tac cgc gtc acc tgg ttc acc tcc tgg agc ccc tgc tac gac tgt gcc      352
258 Tyr Arg Val Thr Trp Phe Thr Ser Trp Ser Pro Cys Tyr Asp Cys Ala
259                  80                  85                  90
261 cga cat gtg gcc gac ttt ctg cga ggg aac ccc aac ctc agt ctg agg      400
262 Arg His Val Ala Asp Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg
263              95                 100                 105
265 atc ttc acc gcg cgc ctc tac ttc tgt gag gac cgc aag gct gag ccc      448
266 Ile Phe Thr Ala Arg Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro
267         110                 115                 120
269 gag ggg ctg cgg cgg ctg cac cgc gcc ggg gtg caa ata gcc atc atg      496
270 Glu Gly Leu Arg Arg Leu His Arg Ala Gly Val Gln Ile Ala Ile Met
271     125                 130                 135
273 acc ttc aaa gat tat ttt tac tgc tgg aat act ttt gta gaa aac cat      544
274 Thr Phe Lys Asp Tyr Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His
275 140                 145                 150                 155
277 gaa aga act ttc aaa gcc tgg gaa ggg ctg cat gaa aat tca gtt cgt      592
278 Glu Arg Thr Phe Lys Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg
279                 160                 165                 170
281 ctc tcc aga cag ctt cgg cgc atc ctt ttg ccc ctg tat gag gtt gat      640
282 Leu Ser Arg Gln Leu Arg Arg Ile Leu Leu Pro Leu Tyr Glu Val Asp
283             175                 180                 185
285 gac tta cga gac gca ttt cgt act ttg gga ctt tgatagcaac ttccaggaat    693
286 Asp Leu Arg Asp Ala Phe Arg Thr Leu Gly Leu
287         190                 195
289 gtcacacacg atgaaatatc tctgctgaag acagtggata aaaaacagtc cttcaagtct    753
290 tctctgtttt tattcttcaa ctctcacttt cttagagttt acagaaaaaa tatttatata    813
291 cgactcttta aaaagatcta tgtcttgaaa atagagaagg aacacaggtc tggccaggga    873
292 cgtgctgcaa ttggtgcagt tttgaatgca acattgtccc ctactgggaa taacagaact    933
293 gcaggacctg ggagcatcct aaagtgtcaa cgtttttcta tgacttttag gtaggatgag    993
294 agcagaaggt agatcctaaa aagcatggtg agaggatcaa atgtttttat atcaacatcc   1053
295 tttattattt gattcatttg agttaacagt ggtgttagtg atagattttt ctattctttt   1113
296 cccttgacgt ttactttcaa gtaacacaaa ctcttccatc aggccatgat ctataggacc   1173
297 tcctaatgag agtatctggg tgattgtgac cccaaaccat ctctccaaag cattaatatc   1233
298 caatcatgcg ctgtatgttt taatcagcag aagcatgttt ttatgtttgt acaaaagaag   1293
299 attgttatgg gtggggatgg aggtatagac catgcatggt caccttcaag ctactttaat   1353
300 aaaggatctt aaaatgggca ggaggactgt gaacaagaca ccctaataat gggttgatgt   1413
301 ctgaagtagc aaatcttctg gaaacgcaaa ctcttttaag gaagtcccta atttagaaac   1473
302 acccacaaac ttcacatatc ataattagca aacaattgga aggaagttgc ttgaatgttg   1533
303 gggagaggaa aatctattgg ctctcgtggg tctcttcatc tcagaaatgc caatcaggtc   1593
304 aaggtttgct acattttgta tgtgtgtgat gcttctccca aaggtatatt aactatataa   1653
305 gagagttgtg acaaaacaga atgataaagc tgcgaaccgt ggcacacgct catagttcta   1713
306 gctgcttggg aggttgagga gggaggatgg cttgaacaca ggtgttcaag gccagcctgg   1773
307 gcaacataac aagatcctgt ctctcaaaaa aaaaaaaaaa aaaaagaaag agagagggcc   1833
308 gggcgtggtg gctcacgcct gtaatcccag cactttggga ggccgagccg ggcggatcac   1893
309 ctgtggtcag gagtttgaga ccagcctggc caacatggca aaaccccgtc tgtactcaaa   1953
310 atgcaaaaat tagccaggcg tggtagcagg cacctgtaat cccagctact tgggaggctg   2013
311 aggcaggaga atcgcttgaa cccaggaggt ggaggttgca gtaagctgag atcgtgccgt   2073
```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 07/19/2002
PATENT APPLICATION: US/09/966,880A TIME: 13:48:24

Input Set : A:\06501-088001.txt
Output Set: N:\CRF3\07192002\I966880A.raw

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; N Pos. 2413

VERIFICATION SUMMARY DATE: 07/19/2002
PATENT APPLICATION: US/09/966,880A TIME: 13:48:24

Input Set : A:\06501-088001.txt
Output Set: N:\CRF3\07192002\I966880A.raw

L:43 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:1
L:46 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:1
L:50 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:1
L:132 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:2366
L:232 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:7
L:235 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:7
L:371 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:9
L:374 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:9